Heavy Precipitation at the Alpine South Side and Saharan Dust over Central Europe: A Predictability Study Using TIGGE

LARS WIEGAND AND ARWEN TWITCHETT

School of Earth and Environment, University of Leeds, Leeds, United Kingdom

CORNELIA SCHWIERZ

Seminar for Statistics, ETH Zurich, Zurich, Switzerland

PETER KNIPPERTZ

School of Earth and Environment, University of Leeds, Leeds, United Kingdom

(Manuscript received 21 December 2010, in final form 22 June 2011)

ABSTRACT

Around 26 May 2008 a pronounced potential vorticity (PV) streamer penetrated from the North Atlantic into the western Mediterranean Sea followed by widespread dust mobilization over the Maghreb region of northwest Africa and a subsequent northward transport into central Europe. At the same time, strong southerly flow over the Mediterranean Sea caused heavy precipitation and flooding at the windward side of the European Alps. Using continuous and feature-based error measures, as well as ensemble correlation techniques, this study investigates the forecast quality and predictability of synoptic and mesoscale aspects of this high-impact event in operational ensemble predictions from nine meteorological centers participating in The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) project. TIGGE is a recently established program providing ensemble forecasts in a standardized format, which allows for an exciting new multimodel approach to investigating the predictability of, for example, high-impact weather and its dynamics. The main conclusions from this study are that 1) the quality of the PV streamer forecasts degrades with lead time showing a general tendency toward too weak Rossby wave; 2) when focusing on the region around the streamer, most models show root-mean-square errors of the same magnitude or larger than the ensemble spread (underdispersive behavior); 3) errors are reduced by about 50% if the comparison is made to each center's own analysis instead of the ECMWF analysis; 4) peak wind speeds over the Sahara tend to be underpredicted, with differences in model formulation dominating over differences in the representation of the PV streamer; and 5) ensemble-mean multimodel forecasts of 4-day accumulated precipitation appear accurate enough for a successful severe-weather warning.

1. Introduction

Rossby wave trains are the dominant dynamical feature of the upper troposphere in the midlatitudes (e.g., Hoskins and Ambrizzi 1993). They propagate eastward with the mean flow and their associated ridges and troughs are related to high and low pressure systems at the surface. In cases of nonlinear amplification, Rossby wave breaking (RWB) can occur. The regions with the highest occurrence of RWB are at the downstream end of the storm tracks (Wernli and Sprenger 2007). Berggren et al. (1949) (and reproduced by Rossby 1959) were among the first to show the propagation of a Rossby wave train over the North Atlantic, its amplification, and the nonlinear wave breaking. The fastest Rossby wave propagation is connected to the jet stream, which acts as a waveguide (Schwierz et al. 2004) and coincides with the strongest isentropic gradients of potential vorticity (PV) that characterize the extratropical tropopause (Hoskins et al. 1985). RWB can result in elongated tongues of high-PV stratospheric air extending equatorward and

Corresponding author address: Lars Wiegand, School of Earth and Environment, University of Leeds, LS92JT Leeds, United Kingdom.
E-mail: l.wiegand@see.leeds.ac.uk

downward into the troposphere. Finally, narrow filaments called PV streamers can develop and eventually break up into distinct PV cutoffs (Appenzeller and Davies 1992). Using idealized modeling, Davies et al. (1991) and Thorncroft et al. (1993) demonstrated that stratospheric filaments can form through dry dynamics and are part of the baroclinic life cycle. Massacand et al. (2001) discuss the idea that upstream diabatic heating might be crucial for the development of some real-world PV streamers.

PV streamers are often connected to high-impact weather events. Poleward moisture transport east of PV streamers can lead to heavy precipitation and flooding (e.g., Tripoli et al. 2005; Knippertz and Martin 2005, 2007a). Studies centered on the European Alps (Massacand et al. 1998; Fehlmann et al. 2000; Hoinka and Davies 2007) have concluded that heavy precipitation events are closely connected with elongated stratospheric intrusions over western Europe, although not all PV streamers produce such events. Martius et al. (2006) found a link between longer-lived PV streamers and greater amounts of accumulated precipitation as well as a higher probability of intense precipitation events. At low latitudes, PV streamers can trigger tropical convection (Kiladis and Weickmann 1992; Slingo 1998), as well as mid- and high-level poleward- and eastward-extending cloud bands from the tropics into the subtropics denoted as tropical plumes (McGuirk et al. 1987, 1988; Iskenderian 1995; Knippertz 2005). Large-scale Saharan dust outbreaks are another significant meteorological phenomenon associated with PV streamers penetrating into the subtropics (Barkan et al. 2005; Slingo et al. 2006; Knippertz and Fink 2006). Thorncroft and Flocas (1997) stated that low-level baroclinicity over the Sahara and low-latitude cyclogenesis producing a cold front along with high winds and low visibility originates from PV anomalies arising "naturally" during the cyclonic wave-breaking process. PV streamers west of the Alps extending into Africa can initiate dust emissions and the transport and deposition on Alpine glaciers (Sodemann et al. 2006). Forecasts of PV streamers, as weather forecasts in general, are limited by deficiencies in numerical weather prediction models and uncertainties in the initial conditions. Ensemble forecasts provide probability information and are therefore of more value for assessing such limitations than deterministic forecasts alone (Tracton and Kalnay 1993). To represent initial condition uncertainties, different methods are used. The European Centre for Medium-Range Weather Forecasts (ECMWF) uses singular vectors to achieve maximum perturbation growth for a given (48 h) optimization time (Palmer et al. 1992). The National Centers for Environmental Prediction (NCEP) use the ensemble transform technique (Wei et al. 2008), which is an improved version of the breeding vector technique (Toth and Kalnay 1993). The breeding perturbations use previous ensemble forecasts to obtain the growing components of the analysis error. As numerical weather prediction models cannot resolve many small-scale features in the atmosphere, these have to be parameterized. In 1998 a stochastic physics scheme was implemented at the ECMWF to represent errors due to parameterizations (Buizza et al. 1998).

So far, comparisons between ensembles from different centers are rare. Buizza et al. (2005) compared the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS), National Centers for Environmental Prediction (NCEP) EPS, and the Canadian Meteorological Centre (CMC) EPS in a comprehensive study. Bourke et al. (2004) tested ECMWF and Australian Bureau of Meteorology (BoM) predictions on their performance in the Southern Hemisphere. Both studies find a better level performance for the ECMWF model and suggest a relation to the superior model and data assimilation system. Other comparisons between the NCEP EPS and the ECMWF EPS have been performed by Atger (1999), Wei and Toth (2003), and Froude et al. (2007).

In the past, comparisons between different ensemble systems were complicated by restricted data availability and the lack of a standardized format. Therefore, the World Meteorological Organization's (WMO) The Observing System Research and Predictability Experiment (THORPEX, information online at http://www.wmo.int/pages/prog/arep/wwrp/new/thorpex_new.html) established a new data archive called the THORPEX International Grand Global Ensemble (TIGGE) that comprises ensemble predictions from different meteorological centers. Daily updated TIGGE forecast products and verifications are provided by Matsueda and Nakazawa online (http://tparc.mri-jma.go.jp/TIGGE/index.html). TIGGE was initiated at a workshop at the ECMWF in 2005 (Richardson et al. 2005) to enhance collaborations in the development of EPSs between operational centers and universities by increasing the availability of data for research. Since 1 February 2008, 10 operational weather forecasting centers have been delivering near-real-time (48-h delay) ensemble forecast data in the same format as is used in three TIGGE data archives, located at the ECMWF, the National Center for Atmospheric Research (NCAR), and the Chinese Meteorological Agency (CMA). First studies using the TIGGE database were published in the last 3 yr. Pappenberger et al. (2008) proposed using TIGGE for early flood forecasting and warning. Park et al. (2008) analyzed the levels of forecast performance of the different

models in the upper troposphere. Matsueda (2009) studied the predictability of atmospheric blocking events. Froude (2010) concentrated on the prediction of extratropical cyclone tracks. Johnson and Swinbank (2009) investigated a multimodel ensemble combining ECMWF, NCEP, and Met Office (UKMO) EPS results and found it to be superior to a single-model ensemble, mainly with respect to surface temperature, but also for geopotential height at 500 hPa and mean sea level pressure. These examples show the wide range of possible research activities and applications feasible with forecasts from the TIGGE database.

The present study adds several new aspects to the published TIGGE research. Using novel feature-based and continuous analysis tools, it thoroughly assesses the forecast performance and predictability of a pronounced RWB event and some of its impacts such as strong near-surface winds and heavy precipitation. Interensemble comparisons reveal both model errors and problems with the design of the EPS as well as the added value of a multimodel ensemble. Furthermore, the study examines the dynamical linkage between the forecast qualities of upper- and lower-level features with a novel correlation approach.

The paper continues with the data and methodology section. Section 3 describes the synoptic evolution of the PV streamer under study and its impacts. In section 4 the forecast performance with respect to the large-scale evolution is presented. In section 5, two types of related high-impact weather—a dust storm as well as heavy precipitation—are evaluated. Conclusions are drawn in section 6.

2. Data and methodology

a. Data

The main database for this study is the TIGGE archive. The 10 participating weather centers in TIGGE are: BoM, CMA, CMC, the Brazilian Centre for Weather Prediction and Climate Studies (Centro de Previsão de Tempoe Estudos Climáticos, CPTEC), ECMWF, the Japanese Meteorological Agency (JMA), the Korea Meteorological Administration (KMA), NCEP, UKMO, and Météo-France. The focus here is on 1–10-day forecasts. Only 1200 UTC forecasts were used for this study because 9 out of 10 centers start forecasts at this time. Unfortunately, Météo-France could not be included in this study because the maximum available forecast lead time is 2.5 days and forecasts start at 1800 UTC. The models of the different centers differ in many aspects. First, the number of ensemble members ranges from 14 to 50 and the number of vertical levels from 19 to 62. Second, the ensemble systems use different data assimilation

techniques as well as different methods of perturbing the initial conditions. Only some models add perturbations to the model physics. Third, the equivalent horizontal resolution of the forecast models range from 0.5° to 1.5°. However, only two models have coarser resolutions than 1°. All meteorological fields in this study are extrapolated onto a 1° × 1° grid. More details on TIGGE are available in, for example, Bougeault et al. (2010) and Froude (2010, her Table 1).

For the verification of the different ensemble prediction systems, analysis data from each TIGGE center and additionally several observational products were considered. These include the gridded precipitation data from the ENSEMBLES E-OBS dataset (Haylock et al. 2008), and a satellite product based on brightness temperature differences from three infrared channels measured by the Spinning Enhanced Visible and Infrared Imager (SEVIRI) instrument flying on the geostationary Meteosat Second Generation (MSG) satellites (for details, see, e.g., Schepanski et al. 2007).

b. Methodology

To study the forecast performance of the different models, vertically averaged PV is considered (Hoskins 1983). On the synoptic scale the full Ertel PV is usually well approximated by the product of the vertical stability and absolute vertical vorticity in pressure coordinates:

$$\text{PV}_p \cong -g\frac{\partial\Theta}{\partial p}(f + \xi), \tag{1}$$

where f is the Coriolis parameter, ξ the relative vorticity, Θ the potential temperature, and g the gravitational acceleration. Both Θ and ξ are either output directly into TIGGE models or can be easily derived from model variables.

Since the TIGGE archive contains only selected pressure levels (200, 250, 300, and 500 hPa), the vertical derivatives in Eq. (1) have to be approximated. Therefore, the PV is reformulated using finite differences. Equation (2) is used for calculating the upper-tropospheric PV between neighboring levels (e.g., 200 and 250 hPa, 250 and 300 hPa, 300 and 500 hPa):

$$\begin{aligned}\text{PV}_p &\cong -g\frac{\partial\Theta}{\partial p}(f + \bar{\xi})\\ &\cong -g\frac{\Theta_{p_{\text{lower}}} - \Theta_{p_{\text{upper}}}}{\Delta p}\left(f + \frac{\xi_{p_{\text{lower}}} + \xi_{p_{\text{upper}}}}{2}\right), \tag{2}\end{aligned}$$

where Δp is the pressure thickness of the layers. The final vertically averaged upper-level PV (UPV) value is computed as the weighted average over the resulting PV values of the three layers. The weights are determined by the thickness of the layers in pressure

TABLE 1. Numbers of EPS members that do not match the analyzed PV streamer. For the details on the method, see section 2b.

Model	Members	24 h	48 h	72 h	96 h	120 h	144 h	168 h
BoM	32	0	0	0	0	12	1	2
CMA	14	0	0	0	3	3	2	3
CMC	20	1	7	7	5	7	10	12
CPTEC	14	0	0	0	0	0	1	0
ECMWF	50	5	10	5	7	4	4	22
JMA	50	0	7	2	8	14	10	12
KMA	16	0	0	6	5	5	2	4
NCEP	20	0	0	6	4	6	5	9
UKMO	23	0	1	2	2	10	8	6

coordinates. A comparison between this UPV approximation and one using all ECMWF model levels shows similar structures but somewhat smaller absolute values (not shown).

Two complementary error measures were applied to the different ensemble forecasts for this case. First, UPV spatially averaged within a box surrounding the PV streamer was considered. Comparisons between ensemble means and analysis fields were undertaken using the root-mean-square error (RMSE) and ensemble spread (standard deviation of the EPS members) for the different operational centers. The differences of these two measures indicate whether the ensemble forecasts are under- or overdispersive for this synoptic situation. Second, a feature-based error measure was applied to the PV streamer. The error assessment compares structural differences within the forecast and analyzed streamers.

The PV streamers were identified using the method introduced by Wernli and Sprenger (2007). Given that the UPV values used here are somewhat smaller than standard PVs, the tropopause is defined as the 1.5-PVU contour in contrast to the often-used 2-PVU value [1 PV unit (PVU) = 10^{-6} m^2 s^{-1} K kg^{-1}]. To identify narrow filaments of stratospheric air, the algorithm searches for segments with an across-contour distance of ≤800 km. The along-contour length connecting these points is required to be ≥1500 km. When these conditions are met, the area enclosed by the contour is identified as a PV streamer.

Matches between the forecast and analysis data are obtained by comparing the positions of the PV streamers. If no overlapping points are found in the analysis and forecast, the closest streamer with a maximum distance of 500 km is considered to be a match. In addition, if the distance between the line points, marking the connection to the main body of stratospheric air, is more than 2000 km away from each other, the streamers are considered to be different systems. Cutoffs and larger troughs of stratospheric air are not included in the comparison, which reduces the number of ensemble

members considered for this analysis for longer lead times (cf. Table 1). The center of mass is used to compare the position of the PV streamer in the forecast in relation to the analyzed PV streamer. For the interpretation of the result it should be borne in mind that minor changes in the shape of the stratospheric air mass can substantially change the northernmost point of an identified PV streamer, which in turn has a large impact on the center position and the forecast error.

3. Synoptic overview

On 20 May 2008 a pronounced upper-level disturbance started to penetrate from the midlatitude North Atlantic into the Mediterranean Sea. Figure 1 shows the time evolution of this trough between 21 and 26 May 2008 using UPV. The streamer originates from a PV reservoir over the North Atlantic and northern Europe on 21 May 2008 (Fig. 1a). A surface low pressure pattern coincided with this PV reservoir. In the course of the following days a large-scale ridge forms to the west of the developing PV streamer (Fig. 1b) and extends far northward between 23 and 24 May 2008 (Fig. 1c). At the surface, high pressure builds under the ridge during this period. The PV streamer is situated between two strong ridges, forming a remarkable "double" omega situation. After 24 May 2008 the PV streamer elongates and reaches its southernmost extension on 26 May 2008 (Fig. 1d). At this time, a surface low forms in the lee of the Atlas Mountains to the southeast of the PV streamer.

The associated strong surface winds over the northern Sahara caused widespread dust mobilization and a subsequent transport across the Mediterranean Sea and Italy into central Europe. Figure 2a shows the ECMWF analysis of 925-hPa winds at 1200 UTC 26 May 2008. The black box defines a region of high winds over the Saharan desert with values of up to 18 m s^{-1}. Satellite images during this time period clearly indicate substantial dust emission and transport. As an example, Fig. 2b shows the situation over the Sahara at 2000 UTC 26 May 2008. Substantial dust uplift occurred at the Algerian border with Mali and Mauritania, to the south of the Algerian and Moroccan Atlas Mountains, and over NE Algeria. Some of this dust was transported northward ahead of the PV streamer and caused a significant dust episode in Germany (Klein et al. 2010). In addition, the strong southerly flow associated with the PV streamer produced heavy precipitation and flooding along the windward side of the Alps. Figure 2c shows the 4-day (0600 UTC 26 May–0600 UTC 30 May 2008) accumulated observed precipitation over the Alpine area. According to the Meteonetwork and the Agenzie Regionali per la Protezione Ambientale
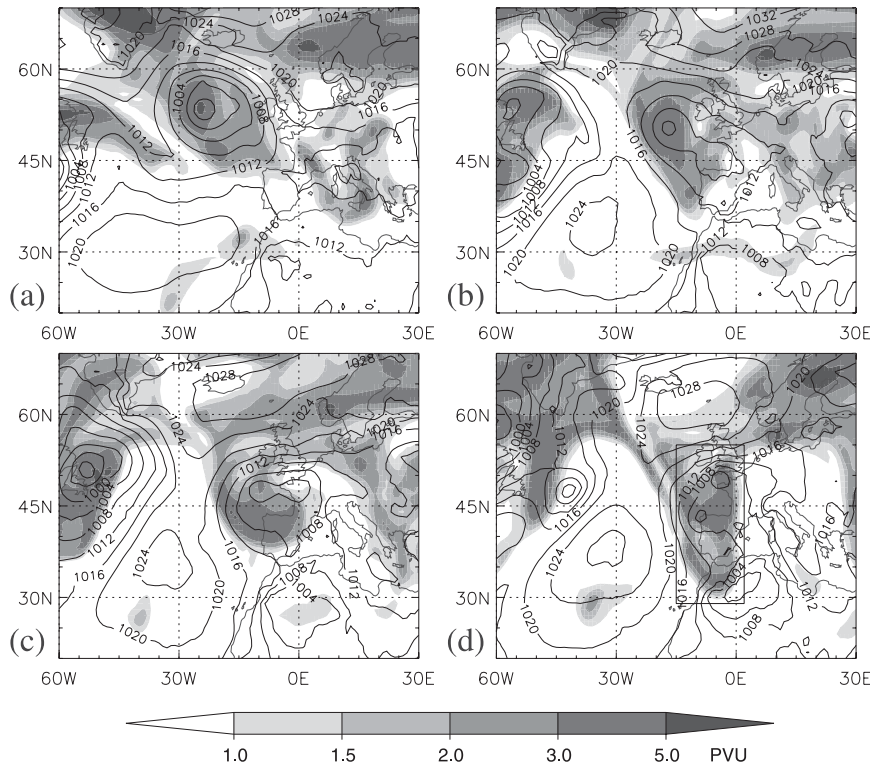
FIG. 1. Synoptic development between 21 and 26 May 2008 as represented by UPV (shaded, see section 2b for definition) and mean sea level pressure (contours every 4 hPa) from ECMWF analysis fields at (a) 1200 UTC 21 May, (b) 1200 UTC 23 May, (c) 1200 UTC 24 May, and (d) 1200 UTC 26 May 2008. The black rectangle in (d) shows the area used for the investigation of the forecast performance.

(ARPA) Piemonte (information available online at http://www.meteo.it/Clima-Statistiche/Analisi-delle28099evento-alluvionale-che-ha-interessato-il-Piemonte-nel-maggio-2008/content/it/1-695-305427-57756) station data, some regions in southern Switzerland and northern Italy recorded daily accumulations of over 100 mm within this time period. The maximum is clearly seen within this area and southeast France in Fig. 2c. The black rectangular box is used for TIGGE forecast performance investigations in section 5c. The heavy precipitation event began on 23 May and the saturated soil received more precipitation starting on 26 May and continuing on 27 and 28 May. According to MeteoSwiss the rapid runoff resulted in flooding on the Alpine south side of Switzerland and Italy, causing Lake Maggiore to rise by 63 cm in just 4 days (information available online at http://www.hydrodaten.admin.ch/d/2022.htm).

4. Forecast performance of large-scale evolution

This section discusses the upper-level development and PV streamer formation in the TIGGE forecast models during this case. A particular focus is on the region of

strong southerly flow from the Sahara across the Mediterranean Sea.

a. Area-mean error measures

Figure 3 shows spaghetti plots of the ensemble-mean 1.5-PVU contours for seven lead times and four selected TIGGE models. In general the PV streamer shortens with increasing lead time in all models with cutoffs occurring in the BoM, ECMWF, and NCEP models. The strongest deviations between lead times are found within the upstream ridge, indicating limited predictability. There is some correlation between the amplitudes of the upstream ridge and the PV streamer when comparing different lead times. Former studies have shown that the western side of high-amplitude ridges over the North Atlantic is often characterized by intense latent heating (Massacand et al. 2001; Knippertz and Martin 2007b). The large spread between models and lead times suggests that the models are sensitive to this process in this situation.

To investigate the forecast performance, two complementary error measures (cf. section 2b) were applied to the TIGGE ensemble forecasts. For RMSE
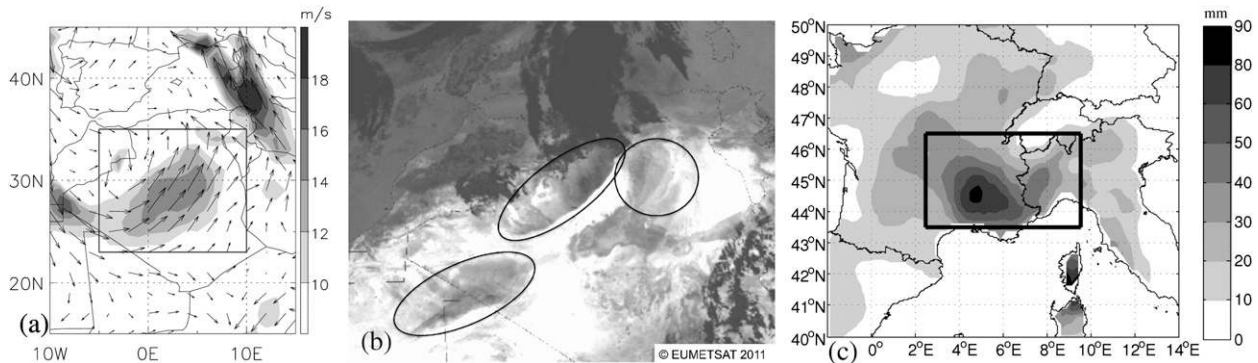
FIG. 2. High-impact weather caused by the PV streamer: (a) ECMWF analysis of wind at 925 hPa at 1200 UTC 26 May 2008. The black rectangle shows the target area used for the investigation of the forecast performance of dust mobilization over the Sahara. (b) Dust uplift over the Sahara represented by the SEVIRI dust product (see section 2b for more information) at 2000 UTC 26 May 2008. Dust stands out as gray areas over the bright desert within the black ellipses. The dark areas to the north are clouds ahead of the upper-level PV streamer. (c) The 4-day precipitation accumulation from 0600 UTC 26 May to 0600 UTC 30 May 2008. The black rectangle shows the target area used for the investigation of the precipitation forecast performance.

computations a target area around the PV streamer under study was defined as 29°–55°N, 15°W–2°E (black box in Fig. 1d). The spatially averaged UPV (grid point average) within this box amounts to 1.45 PVU in the ECMWF analysis. Figure 4 shows results for 1–7-day forecasts of UPV valid at 1200 UTC 26 May 2008 (see Fig. 1d). RMSEs were calculated for ensemble-mean UPV with respect to the ECMWF analysis and each center's own analysis, respectively. Not surprisingly, RMSEs grow with lead time in all models regardless of which analysis is taken. Using the ECMWF analysis as the ''truth'' (Fig. 4a), the ECMWF model clearly shows the best performance for 1–4-day forecasts, and RMSEs of the same magnitude as the other models for longer lead times. When using each center's own analysis instead (Fig. 4b), a substantial decrease in RMSE is apparent for all models. The ECMWF forecast results are still slightly better than those of the other models for short lead times, but for longer lead times ECWMF performs less well. To investigate the variability of the PV fields under consideration, calculations of UPV standard deviations over the box for all models and lead times have been calculated. They result in similar values for all models and a general decrease with increasing lead time (not shown). Therefore, one can exclude the notion that models get penalized by the RMSE calculations (which favor smooth fields) for having more structure in their forecasts.

The increase in RMSE when using ECMWF analysis instead of the center's own analysis amounts to 40% for some models and longer lead times (72–168 h). In the 48-h forecast, RMSEs increase between 6% (NCEP) and 60% (KMA). In the very short-range forecast (24 h),

the increase covers a range for 17% (NCEP) to 75% (JMA, CMA), with an outlier of 150% by the BoM model. The NCEP model shows the smallest RMSE variations between the different calculation methods, indicating that the NCEP and ECMWF analyses are more alike than the others.

To explore this further, differences between all of the TIGGE center's analysis and the ECMWF analysis were calculated. In Fig. 5 the difference between the UKMO and ECMWF analysis is shown as an example. The largest deviations between the two analysis fields are found in the vicinity of the strongest UPV gradients (cf. with Fig. 1d) where values reach up to 1.5 PVU. This is not surprising because small displacements or amplitude errors of PV structures in areas of high gradients can lead to large differences. Comparing short-term forecasts with analysis fields, Dirren et al. (2003) found largest error growth near the jet stream (areas with strongest PV gradients) with some structural similarities to Fig. 5.

To test whether the TIGGE systems are under- or overdispersive, the ensemble-mean RMSE and the spread (standard deviation of all ensemble members) can be compared (Buizza et al. 2005). An ensemble is well designed and reliable, if the RMSE and spread are of similar magnitude at all lead times. Although this concept was proposed for longer time periods of investigation, we apply it here to our case study. Note that the RMSEs are calculated with respect to each center's own analysis. The results for the synoptic situation at 1200 UTC 26 May 2008 are displayed in Figs. 6a–c. For the models CMA, ECMWF, JMA, and KMA, the RMSEs and spreads are similar, with KMA and CMA showing a tendency toward overdispersive behavior at
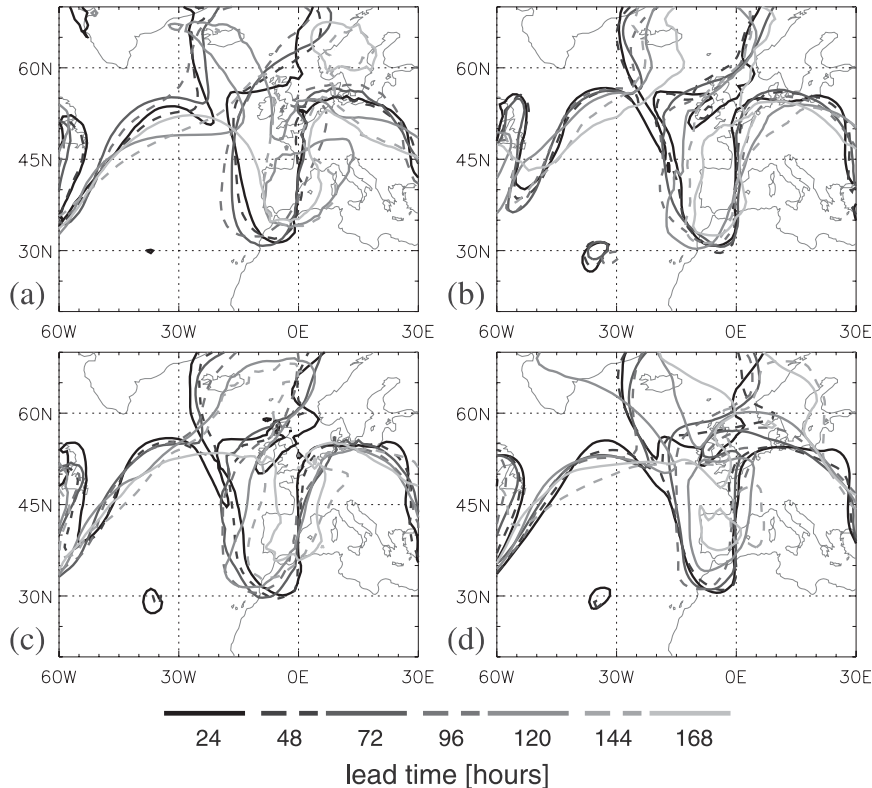
FIG. 3. Forecast ensemble-mean 1.5-PVU contours of UPV for 7 lead times (see legend) and 4 TIGGE centers: (a) BoM, (b) ECMWF, (c) JMA, and (d) NCEP. Forecasts are valid at 1200 UTC 26 May 2008 (see Fig. 1d for comparison).

short lead times (Fig. 6a). BoM, CPTEC, and NCEP exhibit an underdispersive pattern of behavior at all lead times (Fig. 6b). The remaining models CMC and UKMO show changes with lead times (Fig. 6c). CMC is reliable up to a lead time of 96 h but then becomes underdispersive. UKMO is strongly underdispersive except for lead times of 120 and 144 h. A summary view of the model dispersion is presented in Fig. 6d, where the distribution of the differences, RMSE minus spread, is displayed for all seven lead times in the form of box-and-whiskers plots. For a well-designed model, the mean difference should not be significantly different from zero. Although the sample size is rather small and the data are not independent, formal tests (Wilcoxon rank sum test and sign test) have been applied to get a quantitative indication of the model behavior. These tests confirm the visual impression from Figs. 6a–c: significant differences are found for BoM, CPTEC, and NCEP in both tests. For CMC and UKMO the results of the formal tests are inconclusive, probably due to the small sample size and the large variability in the data. In the interpretation of these results the extreme nature of the investigated case has to be kept in mind. Models generally tend to underpredict extremes (Toth 1992;

Ziehmann 2001) and stay closer to climatology, leading to a relatively large RMSE compared to the ensemble spread in such a case.

b. Feature-based error measures

Furthermore, the feature-based error measure explained in section 2b was applied to the PV streamers. It identifies the location and amplitude error for each ensemble member taking the ECMWF analysis as a reference. Figure 7 shows the location error for the nine TIGGE centers, with each lead time represented by a surrounding ellipsoidal envelope. Recall that only matched PV streamers in forecast and analysis are used for this approach. The number of unmatched members per lead time and center is given in Table 1. The center of the ellipse is the mean over all matched ensemble members and the orientation of the main axis is determined by linear regression. The length of this axis is calculated, so that the ellipse encompasses either 90% of the forecast streamer center points or the point with the maximum distance away from the ECMWF analyzed center is reached (which ever occurs first). In this way, the ellipses are sensitive to outliers in order to represent the range of forecasts. The standard deviation of the
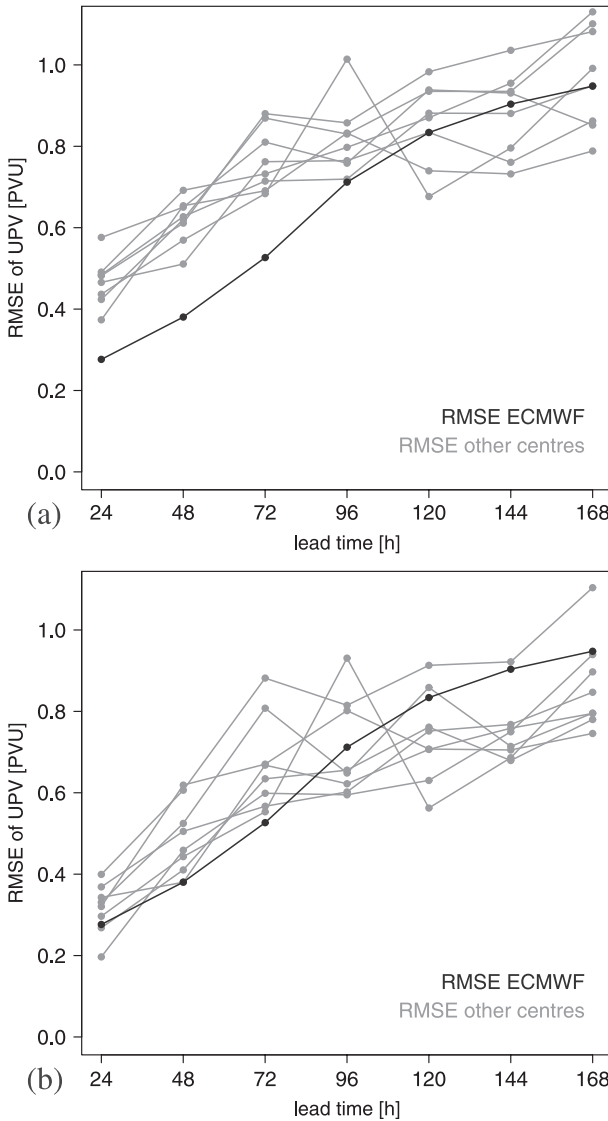
(a)

(b)

FIG. 4. RMSE of ensemble-mean UPV (see section 2b for definition) averaged over the target area (29°–55°N, 15°W–2°E; see Fig. 1d) for each TIGGE center. Forecasts are valid at 1200 UTC 26 May 2008 and have lead times between 24 and 168 h. RMSEs are calculated with respect to (a) ECMWF analyses and (b) each center's own analysis.
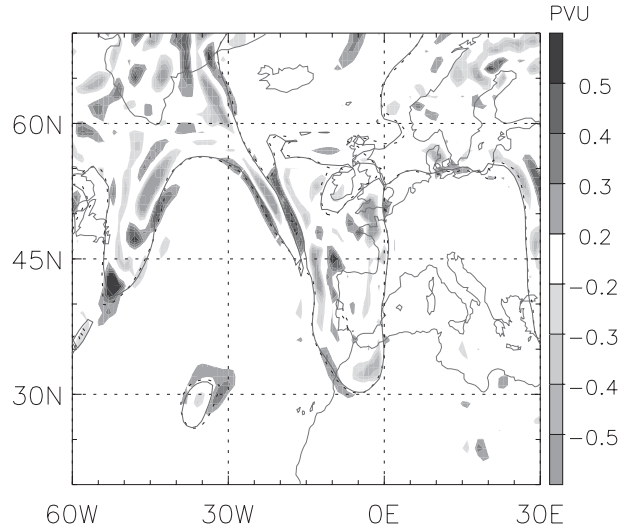

FIG. 5. Difference between UKMO and ECMWF UPV (see section 2b for definition) analysis fields at 1200 UTC 26 May 2008. Positive values indicate larger UKMO UPV. The solid line defines the 1.5-PVU contour line from the ECMWF analysis and the dotted line the 1.5-PVU line of UKMO.

distance between the forecast PV streamer center points and the center point of the ECMWF analysis PV streamer gives the ratio of the major to the minor axis.

The results show a small spread for all centers at short lead times (light gray colors) and a slight northward shift. The spread does not increase substantially over the first four forecasting days and larger spreads are found from the 120-h forecasts onward for all models. This coincides with the increase in the number of nonmatched members in the 120-, 144-, and 168-h forecasts (as shown in Table 1). The nonmatched members are

due to 1) cutoffs; 2) wide stratospheric troughs, which do not fulfill the PV streamer definition (see section 2b); or 3) no stratospheric disturbance at all. The third type of member applies in half of the cases for the 144- and 168-h nonmatches for CMA, CMC, and NCEP, and in one-third of the cases for UKMO. The BoM non-matched members all fall into the first group and the lack of spread in the CPTEC model coincides with few nonmatched cases. With lead times decreasing from 96 to 24 h, most of the members belong to the first group, and the position and size of the cutoffs get closer to that of the streamer. This is also the case for the JMA, KMA, and ECMWF models. The five nonmatched members in the 24-h forecast for ECMWF coincide with cutoffs that are very similar in size and structure to the analyzed PV streamer, with only a thin gap to the main body of stratospheric air. The JMA, KMA, and ECMWF models all show a much wider range of en-semble spread throughout all lead times. The structure of the stratosphere–troposphere interface often differs completely from that in the analysis for the 144- and 168-h forecasts for JMA and ECMWF. The northward shift in most models suggests that the amplitude of the Rossby wave is too small and/or the RWB too weak. The UPV amplitude within the area of the predicted PV streamers determined by the feature-based ap-proach is generally lower than in the ECMWF analysis (not shown). Typical differences range from −0.25 to −0.5 PVU for short lead times and increase to −0.75 to −1.5 PVU for later forecast times.
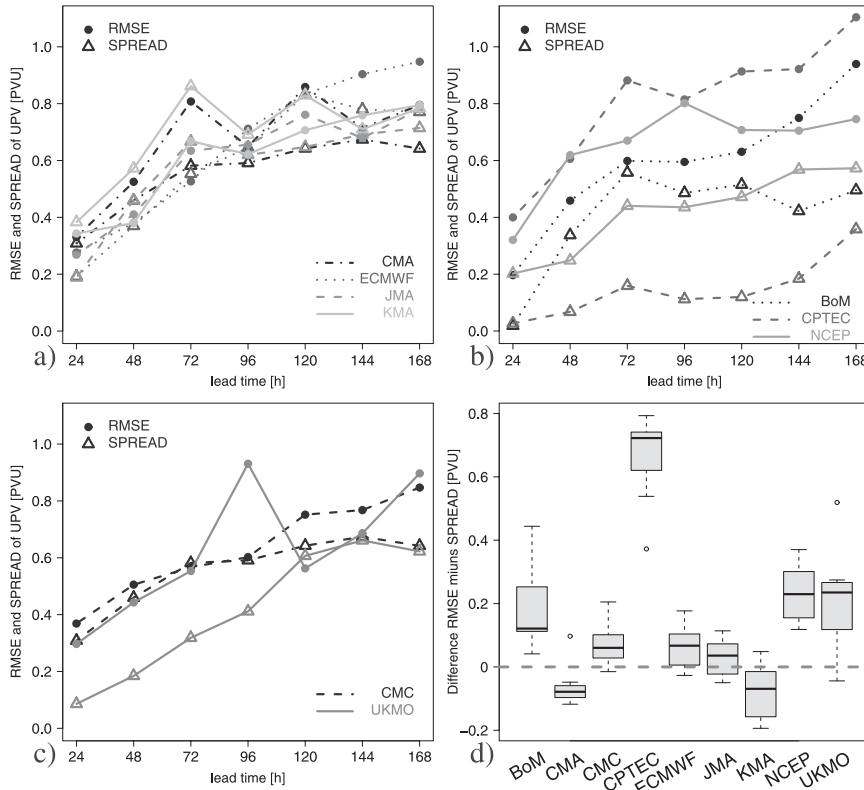
FIG. 6. RMSE as in Fig. 4b (circles) together with the corresponding ensemble SPREAD (standard deviation, open triangles) for (a) nondispersive, (b) underdispersive, and (c) inconclusive (see section 4a for details) behavior of the TIGGE models. (d) Difference between RMSE and SPREAD accumulated over all lead times for each TIGGE center. Forecasts are valid at 1200 UTC 26 May 2008 and have lead times between 24 and 168 h.

5. High-impact weather

Building on the results of the forecast performance for upper levels, this section investigates how TIGGE models forecast the meteorological fields that are important for high-impact weather, in this case dust mobilization and heavy precipitation. The section is divided into four parts. Starting with an analysis of the predicted potential dust uplift (PDU) based on low-level winds over the Sahara, it continues with a correlation between PDU and UPV, followed by an investigation of the predicted precipitation and its relation to UPV forecasts.

a. Dust mobilization

As pointed out in section 3, large amounts of dust were lifted over the northern Sahara and carried into Europe in connection with the PV streamer under study. As the TIGGE models do not provide information on surface stress and soil characteristics, we will use a simple diagnostic parameter that reflects the essence of the physical dependency of dust emission on wind speed

based on the widely used dust uplift parameterization by Marticorena and Bergametti (1995). If we, for the sake of simplicity, assume the soil to be homogeneous with a constant emission threshold and approximate the surface stress or friction velocity by its first-order control low-level wind speed U, we can define the PDU in the following way:

$$\text{PDU} = U^3 \left(1 + \frac{U_t}{U}\right)\left(1 - \frac{U_t^2}{U^2}\right), \qquad (3)$$

where U_t denotes the constant threshold velocity that is typically on the order of 8 m s^{-1} at 10 m above the ground (Chomette et al. 1999). As 10-m winds are not a prognostic variable in TIGGE models, it has to be computed through extrapolation from the lowest model level using parameters from the boundary layer scheme. The details of this procedure can lead to dramatic differences in low-level wind speeds (see, e.g., Schwierz et al. 2010). To avoid that the differences between the models will be dominated by the effects of parameterizations, wind speed at 925 hPa was used as a robust
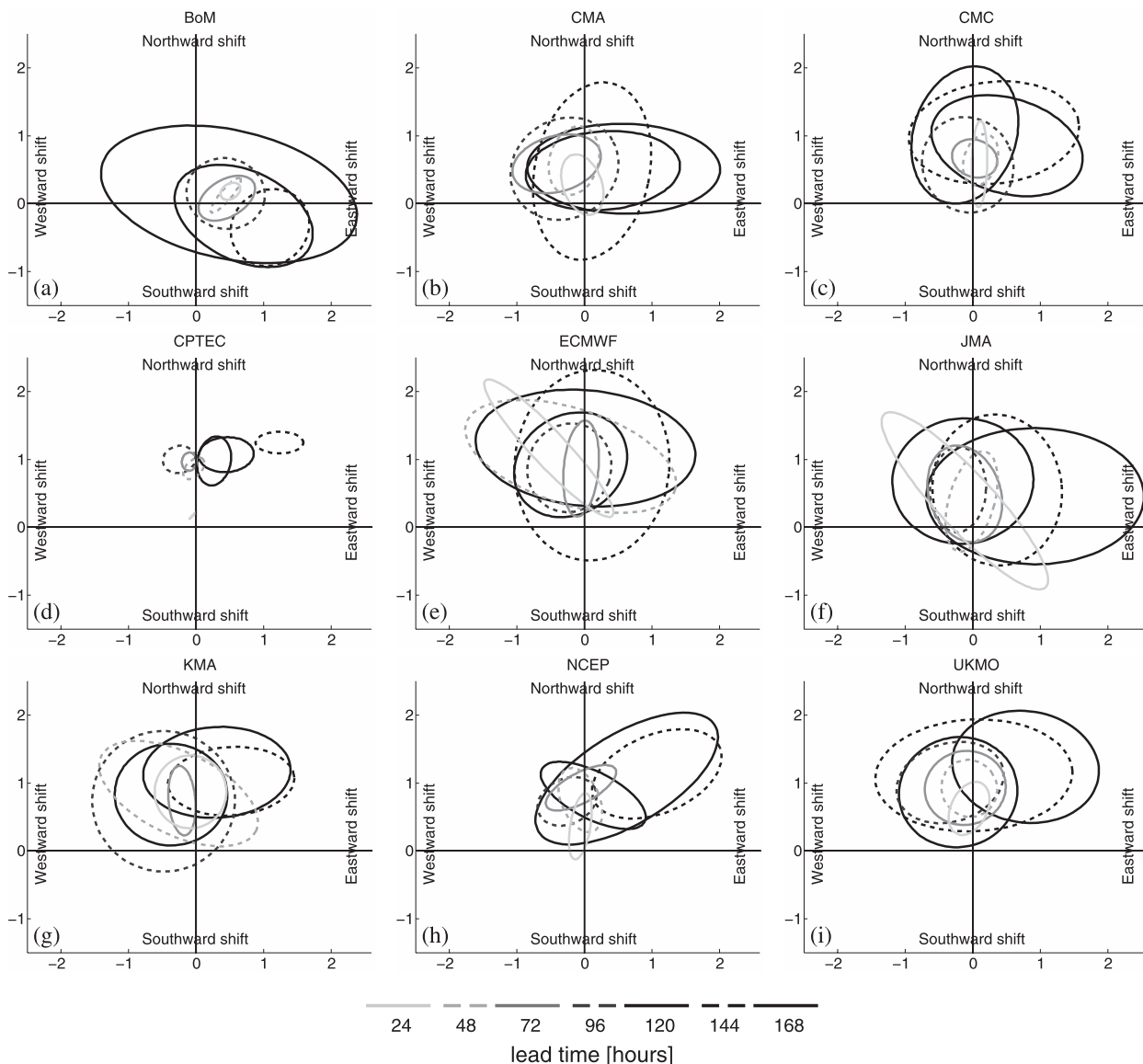
FIG. 7. Position error (in 1000 km) of the streamer ensemble forecasts based on UPV (see section 2b for definition) from 9 TIGGE forecast centers (indicated in each panel). Ellipses surround position forecasts from all members at 1200 UTC 26 May 2008 (see Fig. 1d). Gray shading indicates lead time ranging from 24 to 168 h (see legend). Center of the coordinate system is the PV streamer position as analyzed by ECMWF in all panels.

estimate for the flow in the boundary layer over Africa with U_t set to 10 m s^{-1}. In principle, it would be desirable to evaluate wind direction as well, but for the sake of simplicity the authors decided to concentrate on the main atmospheric control on dust emission (i.e., wind speed).

Figure 8 displays PDU averaged over the target box shown in Fig. 2a for the nine TIGGE ensemble systems in the form of box-and-whiskers plots. The size of the boxes is indicative of the ensemble spread (see caption). The dashed lines show PDU values computed from each

center's analysis. The large range from 0.52×10^3 to 1.43×10^3 m^3 s^{-3} is a reflection of the differences between the individual analyses, as already discussed in section 4a, and of the highly nonlinear relation between dust emission and wind speed (essentially a cubic dependence).

The 24-h forecasts (Fig. 8a) show a number of different patterns of behavior for the ensemble systems. CPTEC and BoM have a very low spread and clearly underpredict PDU, while UKMO shows an equally small spread with a slight overprediction. CMA stands out as the model
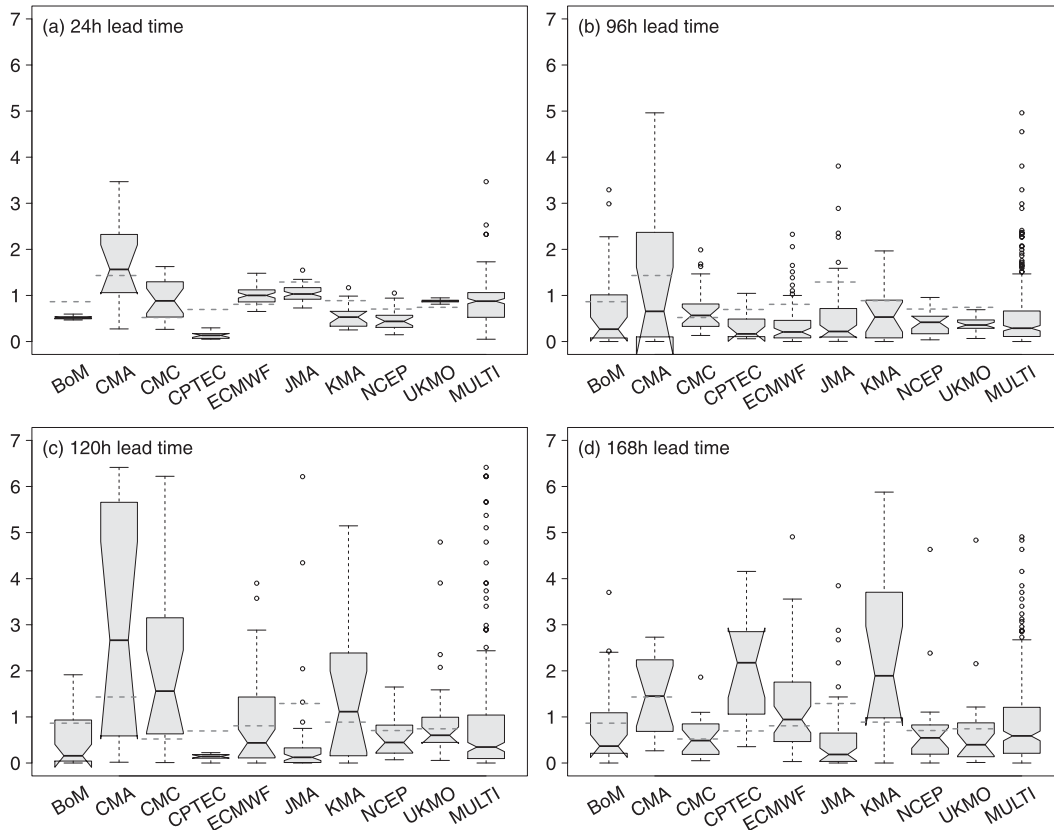
FIG. 8. Potential dust uplift (10^3 m^3 s^{-3} PDU; see section 5a for definition) averaged over the box 23°–35°N, 5°W–10°E from each TIGGE center and the multimodel ensemble including all TIGGE members for (a) 24-, (b) 96-, (c) 120-, and (d) 168-h forecasts at 1200 UTC 26 May 2008. The dashed gray lines indicate the PDU analysis from each TIGGE center. Box-and-whiskers plots are defined with the upper and lower borders of the gray box as the first and third quartiles, respectively; the black line within the box as the median; the whiskers on both sides as the minimum and maximum; and the blank circles as outliers. The notches on the boxes indicate that the medians of two box plots are significantly different when notches do not overlap.

with the largest spread and the highest median and analyzed PDU. CMC also has a large spread, but the lowest analyzed PDU and a median close to most of the other models. Intermediate spreads that contain the analyzed values are found for ECMWF, JMA, KMA, and NCEP; the last threes' median clearly underpredicting PDU over the Sahara. The multimodel box-and-whiskers plot including all TIGGE members has a median exactly in the range of the analyzed values, demonstrating the added value of a multimodel approach. This also holds for 48- and 72-h forecasts (not shown).

The 96-h PDU forecasts (Fig. 8b) show a marked increase in spread and a general tendency to underpredict, with most medians well below the analyzed PDU including the multimodel one. UKMO is the only model where the analyzed value is outside of the (rather small) ensemble spread. For the 120-h PDU predictions (Fig. 8c), the spread continues to increase for most ensemble systems and comprises the analyzed PDU for all models but

CPTEC, which simulates hardly any wind speeds above the emission threshold. Out of the remaining eight models, five show an underprediction. CMA, CMC, and KMA show the largest spreads and tend to overpredict the analyzed PDU. Interestingly, several ensemble systems have members with rather extreme PDUs on the order of 5 times the analysis value.

For the longest lead time of 168 h (Fig. 8d), all models contain the analysis in their ensemble spread, which does not increase much more beyond the 120-h forecasts. Compared to some of the shorter lead times, the numbers of members lying within the range of analyzed PDUs have increased in all models except for those of UKMO (decreased median) and KMA (higher PDU in total). In particular, CMA and CMC show good forecasts, with the median matching their own analyses exactly. Even CPTEC shows the potential to lift dust in contrast to shorter lead times. This is reflected in the multimodel box-and-whiskers plot, too, which has a higher median
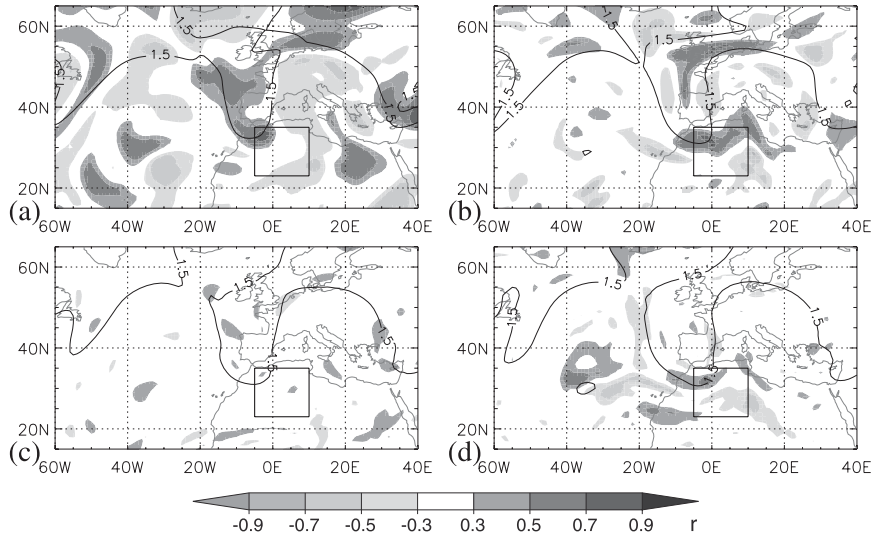
FIG. 9. Correlation coefficients (gray shading) between the predicted potential dust uplift in a box over the Sahara (black rectangle) and the predicted UPV at every grid point, both valid at 1200 UTC 26 May. (a) All lead-time forecasts (24–168 h) from NCEP, (b) 48-h forecast from NCEP, (c) all lead-time forecasts from ECMWF, and (d) 72-h forecast from ECMWF. Black lines are the ensemble mean 1.5-PVU contour line from the respective center at the respective lead time(s).

value for 168 h than for the 96- and 120-h PDU multimodel forecast medians. Given the degradation of the PV streamer forecast with lead time discussed in section 4, it is conceivable that the overall satisfactory quality of the 168-h forecasts could be at least partly due to the wrong physical reasons. This point will be discussed further below (see section 5b).

In conclusion, there is a general tendency in the TIGGE models to underpredict PDU for all lead times, but there are at least some members indicating possible dust mobilization in nearly every forecast. There is a nonmonotonic relationship between forecast quality and lead time with overall large variations in both the median and spread between the modeling systems. The multimodel approach improves the forecast at most lead times. One general problem for the verification is the surprisingly large difference between the model analyses.

b. Ensemble correlation of dust and UPV

This section takes a more detailed look at the relationship between the PDU over the Sahara and the upper-level circulation using an ensemble correlation technique. This method is adapted from Hawblitzel et al. (2007), who used correlations between different forecast variables to investigate precursors for successful forecasts. A linear correlation coefficient r is calculated using

$$r(\mathrm{UPV}_{i,j}^t, \mathrm{PDU}^t) = \frac{\dfrac{1}{N-1}\displaystyle\sum_{n=1}^{N}(^n\mathrm{UPV}_{i,j}^t - \overline{\mathrm{UPV}_{i,j}^t})(^n\mathrm{PDU}^t - \overline{\mathrm{PDU}^t})}{\left[\dfrac{1}{N-1}\displaystyle\sum_{n=1}^{N}(^n\mathrm{UPV}_{i,j}^t - \overline{\mathrm{UPV}_{i,j}^t})^2\right]^{1/2}\left[\dfrac{1}{N-1}\displaystyle\sum_{n=1}^{N}(^n\mathrm{PDU}^t - \overline{\mathrm{PDU}^t})^2\right]^{1/2}}, \tag{4}$$

where PDU^t is the box average (see Fig. 2a) of the potential dust uplift as defined in section 5a at lead time t and N is the number of EPS members (different for every center). The PDU values are correlated with UPV at every grid point (i, j) predicted for the same lead time t. Four exemplary results are shown in Fig. 9 where shadings indicate positive (negative) correlation in dark gray (light gray). A positive (negative) value of r at point (i, j) implies that PDU in the box over the Sahara and UPV values at point (i, j) tend to deviate in the same (opposite) direction from the ensemble mean. The ensemble-mean 1.5-PVU line is included for reference. The top (bottom) panel in Fig. 9 displays results from NCEP (ECMWF) forecasts. The left column in Fig. 9 shows

correlations over all seven lead times ($N = 7 \times$ number of EPS members) and the right column shows one lead time only (48 h for NCEP, and 72 h for ECMWF).

The results for all NCEP forecasts (Fig. 9a) show a strong correlation within the ensemble-mean PV streamer and to the southeast of its tip. Negative correlations are evident in the upstream and downstream ridges. This suggests that the PDU over the Sahara is closely related to the amplitude of the Rossby wave and the southward extension of the PV streamer. Together with the negative correlations over the Sahara, this leads to a stronger PV gradient across the box, which induces a stronger southwesterly flow. Looking at individual members with particularly high and low PDU values (not shown) indicates that the former are characterized by a thinner PV streamer than analyzed, reaching 30° latitude with an enhanced ridge upstream, while the latter tend to show UPV cutoffs and a relatively weak ridge upstream. Similar results are found for UKMO, KMA, and BoM, although the negative correlations to the west of the PV streamer are smaller in these models.

An example for a single lead time is given in Fig. 9b. Generally, the correlation signals show more finescale structure and are somewhat more concentrated around the PV streamer than in Fig. 9a. Indications for a thinning and southward extension are very clear for this lead time with positive values extending well across northern Algeria. This pattern together with negative values over Spain and France occurs for many models and lead times, and is indicative of a so-called treble clef structure resulting from cyclonic role up and the diabatic reduction of PV (Appenzeller et al. 1996; Martin 2006).

Analogous analyses for other TIGGE models indicate a less clear connection between UPV and PDU. For all lead times combined, correlations are generally weak for ECMWF (Fig. 9c), while the JMA, CMA, and CMC models show at least weakly positive correlations within the PV streamer, and the CPTEC model shows negative correlations in the area of the upstream ridge (not shown). For individual lead times, however, some coherent signals can be identified for the ECMWF model such as an enhanced UPV gradient across the box (Fig. 9d). Looking at individual forecasts for these TIGGE models (not shown) confirms that similar PV streamer forecasts can produce both low and high PDU values over the Sahara. This indicates that other processes such as vertical stability modify the physical link between the intensity of the streamer and the low-level winds. A detailed investigation of this interesting aspect is beyond the scope of this paper, but should be borne in mind in future application of ensemble correlation techniques.
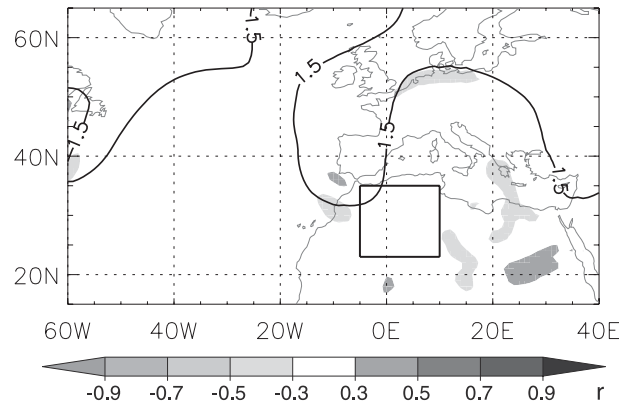

FIG. 10. As in Fig. 9, but for 72-h forecasts from all TIGGE members.

A new research approach made possible through TIGGE is the calculation of multimodel statistics, which allows us to distinguish between inter- and intra-ensemble spreads. Here, ensemble correlation techniques are applied to all 239 TIGGE members [N in Eq. (4)] for every forecast time step. Figure 10 shows the 72-h lead time correlation of UPV and PDU as in Fig. 9d but for all TIGGE forecasts. Overall, the correlations are much weaker than for most single-model, single-lead time calculations. This suggests that differences in model formulation have a stronger effect on the low-level wind speed than differences in UPV between different members. Most likely, different boundary layer formulations play an important role in generating this result.

c. Precipitation

A similar analysis to section 5a is conducted for the heavy rainfall along the southern side of the Alps based on precipitation accumulated over the period from 0600 UTC 26 May to 0600 UTC 30 May and averaged over the box from 43.5°–46.5°N and 2.5°–9.5°E, where the maximum precipitation occurred (see Fig. 2c). After the southernmost extension of the PV streamer at 26 May (see Fig. 1), the PV streamer thinned and shifted eastward until 28 May. Subsequently, it rolled up cyclonically resulting in a PV cutoff over the Alps on 30 May 2008. The cutoff was then approached by a new PV streamer from the northwest, which however did not reach the large southern extension of the previous streamer.

In analogy to Fig. 8, Fig. 11 shows box-and-whiskers plots of predicted precipitation and the observed 4-day box average of 37 mm as a dashed line. This value is much smaller than the maximum values discussed in section 3 due to area averaging during the gridding process of the ENSEMBLES data (see details in Haylock et al. 2008) and large spatial inhomogeneities. Global models cannot be expected to capture this mesoscale spatial variability.
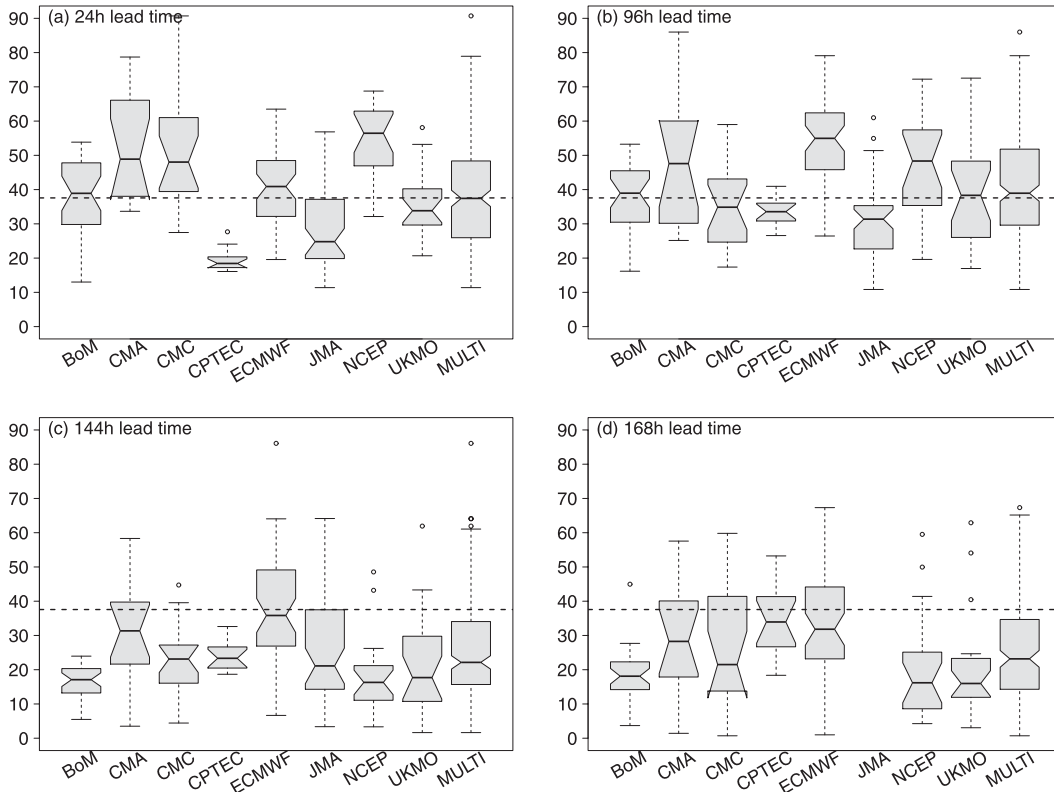
FIG. 11. Predicted precipitation (mm) accumulated over the 4-day period covering 0600 UTC 26 May–0600 UTC 30 May 2008 and averaged over the target box (43.5°–46.5°N, 2.5°–9.5°E) from each TIGGE center and the multi-model ensemble except for KMA (see section 5c for details). Lead times are (a) 24, (b) 48, (c) 96, and (d) 168 h. The dashed line is the observed value from the ENSEMBLES E-OBS dataset. Box-and-whiskers plots are defined as in Fig. 8.

Therefore, the forecasts are more useful as indicators of a potentially significant rain event than for flood warnings on a local scale. In Fig. 11a the lead time of 24 h refers to forecasts initialized at 1200 UTC 25 May, so that the 4-day accumulation is between forecast steps 18 and 114 h. KMA had to be excluded in this analysis, as it does not provide precipitation data. JMA could not be investigated for all lead times because of a lack of precipitation forecasts longer than 240 h.

The 24- and 48-h forecasts (Fig. 11b, initialized at 1200 UTC 24 May, accumulated between forecast times of 42 and 138 h) are very similar and show a relatively large spread but a good amount of scatter around the observed value. The most skillful short-range forecasts are provided by BOM, ECMWF, and UKMO with the median close to the observed rainfall. CMA, CMC, and NCEP slightly overpredict the average rainfall while JMA has lower rainfall averages than those observed. CPTEC has a very low level of spread and significantly underpredicts the rainfall. The forecasts from this model improve over time (although still underpredicting the average rainfall) up to the 96-h forecast (cf. Fig. 11b).

The multimodel box-and-whiskers plot including all TIGGE member forecasts for this lead time shows a remarkably good level of performance with its median exactly on the observed 37 mm of rainfall. The same multimodel pattern of behavior can be seen across the whole short range up to 96 h. The 72- and 96-h forecasts are again similar for each model center with several of the models showing skillful predictions. BoM and UKMO perform best with regard to the median rainfall.

A view at the longer-range forecasts (120–168 h) shows a small overprediction by CPTEC and ECMWF (120 h), while the BoM model notably changes in skill and underpredicts the average rainfall considerably (120–168 h). In the 144-h forecast (Fig. 11c) the spread decreases somewhat in contrast to the increasing spread in UPV (see section 4a). ECMWF performs best for this lead time but all the centers' medians are below the observed precipitation as shown by the multimodel result. This continues for the 168-h forecast (Fig. 11d) with all centers underpredicting the precipitation and BoM, NCEP, and UKMO having the lowest and CMA, CPTEC, and ECMWF the highest medians.
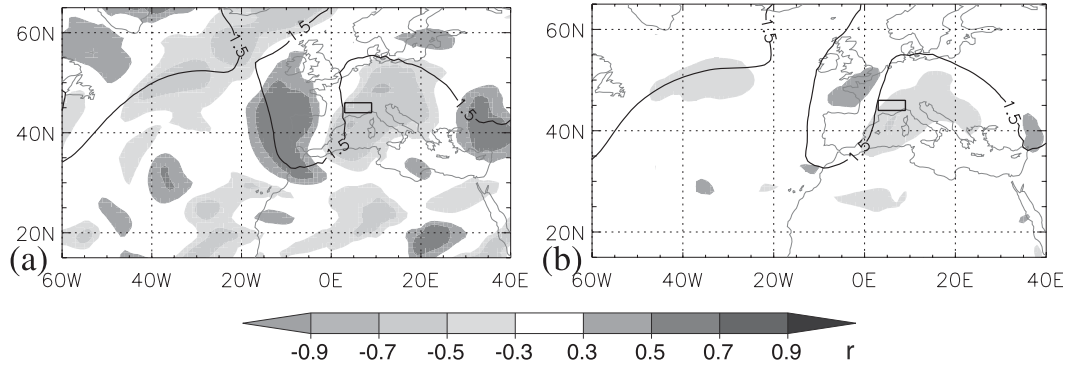
FIG. 12. Correlation coefficients (gray shading) between the predicted 4-day accumulation of precipitation valid from 0600 UTC 26 May to 0600 UTC 30 May in a box over the Alpine south side (black rectangle) and the predicted UPV at every grid point valid at 1200 UTC 26 May for (a) all lead-time forecasts (24–168 h) from BoM and (b) the 120-h forecasts from all TIGGE members. Black lines are the ensemble-mean 1.5-PVU contour lines over (a) all lead times and (b) all TIGGE members.

The most notable changes with lead time occur for the UKMO and BoM models, which show good skill for forecast times up to 4 days and significant decreases beyond this. ECMWF has the most consistent forecasts with accurate short-range (24–72 h) and medium-range prediction (144–168 h). The 24–96-h multimode forecasts give an overall satisfactory representation of the event and could be used for severe weather warnings. Similarly, exceptionally good predictions of heavy precipitation events on the Alpine south side in general were shown, for instance, in Jenkner et al. (2008) for the Consortium for Small-Scale Modeling (COSMO) mesoscale model.

d. Ensemble correlation of precipitation and UPV

The correlation technique explained in section 5b has also been applied to the UPV and box-averaged 4-day accumulated precipitation (0600 UTC 26 May–0600 UTC 30 May 2008). Lead time for UPV is defined with respect to 1200 UTC 26 May 2008. Correlations were also calculated for UPV in the middle of the precipitation period (i.e., 28 May 2008), but the results are quite similar with even smaller signals in some areas. Single-model and single-time-step correlations (not shown) resulted in overall lower values compared to the PDU and UPV correlations shown in Figs. 9b and 9d. This might be caused to some extent by the relatively long accumulation period used for precipitation. A feature evident for most models and lead times is a dipole of positive correlations along the western side of the PV streamer and negative correlations farther east. This highlights the importance of high PV in the western vicinity (France and northern Spain) of the Alpine south side for producing a high amount of precipitation in the following days (cf. also Martius et al. 2006).

As an example, Fig. 12a shows correlations over all lead times for the BoM model. The results exhibit negative correlations within the downstream and upstream ridge and positive values within the PV streamer itself, particularly on the western side. Some indications for cyclonic rollup are evident in the shape of the correlation pattern over northwestern Africa. The dipole indicates that fast-moving PV streamers (high values over Italy and central Europe and low values over the Iberian Peninsula) are associated with too low precipitation over the Alps and vice versa. Looking at the driest and wettest individual forecasts confirms this (not shown). Dry members are also more oriented in a northeast–southwest direction. Again, the negative correlations in the upstream ridge are most likely a reflection of the role of diabatic processes for the amplification of the wave and the subsequent breaking (cf. Fig. 3), while negative values downstream indicate a more effective blocking that slows the propagation of the PV streamer.

For most other models, the correlations over all lead times are small, indicating the presence of different UPV–precipitation relationships that cancel out when all lead times are combined. Correlations for all models at individual lead times are also rather small, as found for PDU and UPV in section 5b, pointing to substantial model differences. The sole exception being at 120 h (Fig. 12b), which shows some structural resemblance to the results for BoM for all lead times (Fig. 12a), albeit with overall smaller values.

6. Summary and conclusions

This paper has analyzed forecasts from nine ensemble prediction systems archived in the TIGGE database for the case of a pronounced PV streamer over the Mediterranean Sea that produced a combination of severe

weather events including a Saharan dust storm and flooding in the southwestern European Alps. Different analysis techniques including RMSE and spread comparisons, differences between analysis fields, and a feature-based verification algorithm were used to investigate the forecast quality and predictability of the PV streamer and near-surface wind as well as precipitation. The connection between upper-level features and low-level impacts was investigated via an ensemble correlation technique.

RMSE and ensemble spread calculations based on upper-level PV in the area of the streamer show increasing values with lead time with marked differences depending on the analysis used for reference. Using the ECMWF analysis instead of every center's own analysis increased RMSEs by up to 75% (60%) in the 24-h (48 h) forecasts. Even in the long-range forecasts (72–168 h) RMSE differences up to 40% occur. Comparisons of analysis fields with each other showed differences of up to 1.5 PVU, mainly in regions of large PV gradients along the midlatitude jet.

The comparison of RMSE and ensemble spread revealed a tendency for under- to nondispersive behavior by most ensemble systems for this rather extreme case. Results from a feature-based measure that focused on PV streamer location errors indicate a general tendency for a northward shift of the PV streamer in the forecasts. This shift and the ensemble spread for every TIGGE center are relatively small until 96-h lead time; they then increase abruptly indicating a limit of predictability beyond this range for this event. The consistent behavior of all TIGGE models suggests general problems with reproducing the intensity of the wave breaking, resulting in too small wave amplitudes and an insufficient meridional extension of the PV streamer to the south.

Using a new potential dust uplift diagnostic for the Saharan dust storm showed a general tendency of underpredicting peak low-level winds, but a large spread in most ensembles gives at least some indication of the potential for dust mobilization. Ensemble spread and forecast quality are nonmonotonic with lead time in contrast to the upper-level features.

Using an ensemble correlation technique, the relationship between predicted upper-level PV fields and potential dust uplift was investigated. Ensemble correlations show that a high-amplitude Rossby wave and a strong PV gradient over the Sahara are important ingredients for low-level winds in that region. Comparing different lead times for single models or different models for single lead times, points to the importance of other factors modifying the relationship between UPV and low-level wind such as vertical stability and boundary layer parameterizations.

Overall, the TIGGE models yield good 4-day accumulated precipitation forecasts at short lead times (1–4 days) with decreasing skill for longer lead times except for ECMWF. Correlation results showed the importance of a well-forecast PV streamer in time and east–west position to get the right amounts of precipitation along the Alpine south side. Both the up- and downstream ridges are important for this. For both parameters (PDU and precipitation), a multimodel approach proves beneficial for most lead times.

This paper is among the first to use the exciting multimodel ensemble capabilities of TIGGE for a detailed predictability case study of a pronounced upper-level PV streamer and its near-surface impacts. Previous studies investigated forecast performance with respect to upper-level development (Park et al. 2008) and cyclone tracks (Froude 2010) in a more climatological sense. Some of the analysis techniques presented here could be applied to more PV streamer cases and other phenomena in future work. When doing this, it is important to keep some of the identified limitations in mind. One caveat that became evident through this study is the limited availability of vertical levels that only allows us to calculate a rather crude upper-level PV. The PV differences used for forecast evaluation and comparing TIGGE analyses can be misleading, as low PV values in the troposphere (0–2 PVU) are compared to the high stratospheric values (2–20 PVU) across the tropopause. Some first thoughts of using a rescaled-PV formulation instead are discussed in Martius et al. (2010). The surprisingly large discrepancies between different analyses strongly suggest that uncertainties in what we regard as the "truth" have to be taken into account much more than is currently done. Froude (2010) for example used only ECMWF analysis data in her TIGGE study and speculated that this could bias the results in ECMWF's favor. This is corroborated by the results of the case study presented here.

TIGGE provides a great opportunity to explore this aspect further in the future. In addition, the comprehensive list of variables that TIGGE includes offers the possibility to focus on dynamical aspects and interaction processes, which can be beneficial for pinpointing some of the predictability issues evident in forecasting high-impact weather events.

(UKMO) for providing access to the ECMWF database. Furthermore, we would like to acknowledge valuable comments from one anonymous reviewer and from Heini Wernli.

REFERENCES

Appenzeller, C., and H. C. Davies, 1992: Structure of stratospheric intrusions into the troposphere. *Nature,* **358,** 570–572.

——, H. Davies, and W. Norton, 1996: Fragmentation of stratospheric intrusions. *J. Geophys. Res.,* **101D,** 1435–1456.

Atger, F., 1999: The skill of ensemble prediction systems. *Mon. Wea. Rev.,* **127,** 1941–1953.

Barkan, J., P. Alpert, H. Kutiel, and P. Kishcha, 2005: Synoptics of dust transportation days from Africa toward Italy and central Europe. *J. Geophys. Res.,* **110D,** D07208, doi:10.1029/2004JD005222.

Berggren, R., B. Bolin, and C.-G. Rossby, 1949: An aerological study of zonal motion, its perturbations and break-down. *Tellus,* **1,** 14–37.

Bougeault, P., and Coauthors, 2010: The Thorpex Interactive Grand Global Ensemble. *Bull. Amer. Meteor. Soc.,* **91,** 1059–1072.

Bourke, W., R. Buizza, and M. Naughton, 2004: Performance of the ECMWF and the BoM ensemble prediction systems in the Southern Hemisphere. *Mon. Wea. Rev.,* **132,** 2338–2357.

Buizza, R., T. Petroliagis, T. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons, and N. Wedi, 1998: Impact of model resolution and ensemble size on the performance of an ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **124B,** 1935–1960.

——, P. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.,* **133,** 1076–1097.

Chomette, O., M. Legrand, and B. Marticorena, 1999: Determination of the wind speed threshold for the emission of desert dust using satellite remote sensing in the thermal infrared. *J. Geophys. Res.,* **104,** 31 207–31 216.

Davies, H. C., C. Schär, and H. Wernli, 1991: The palette of fronts and cyclones within a baroclinic wave development. *J. Atmos. Sci.,* **48,** 1666–1689.

Dirren, S., M. Didone, and H. C. Davies, 2003: Diagnosis of ''forecast–analysis'' differences of a weather prediction system. *Geophys. Res. Lett.,* **30,** 2600, doi:10.1029/2003GL017986.

Fehlmann, R., C. Quadri, and H. C. Davies, 2000: An alpine rainstorm: Sensitivity to the mesoscale upper-level structure. *Wea. Forecasting,* **15,** 4–28.

Froude, L. S. R., 2010: TIGGE: Comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting,* **25,** 819–836.

——, L. Bengtsson, and K. I. Hodges, 2007: The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. *Mon. Wea. Rev.,* **135,** 2545–2567.

Hawblitzel, D. P., F. Zhang, Z. Meng, and C. A. Davis, 2007: Probabilistic evaluation of the dynamics and predictability of the mesoscale convective vortex of 10–13 June 2003. *Mon. Wea. Rev.,* **135,** 1544–1563.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.,* **113D,** D20119, doi:10.1029/2008JD010201.

Hoinka, K. P., and H. C. Davies, 2007: Upper-tropospheric flow features and the Alps: An overview. *Quart. J. Roy. Meteor. Soc.,* **133,** 847–865, doi:10.1002/qj.69.

Hoskins, B. J., 1983: Dynamical processes in the atmosphere and the use of models. *Quart. J. Roy. Meteor. Soc.,* **109,** 1–21.

——, and T. Ambrizzi, 1993: Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.,* **50,** 1661–1671.

——, M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

Iskenderian, H., 1995: A 10-year climatology of Northern Hemisphere tropical cloud plumes and their composite flow patterns. *J. Climate,* **8,** 1630–1637.

Jenkner, J., C. Frei, and C. Schwierz, 2008: Quantile-based short-range QPF evaluation over Switzerland. *Meteor. Z.,* **17,** 827–848, doi:10.1127/0941-2948/2008/0344.

Johnson, C., and R. Swinbank, 2009: Medium-range multimodel ensemble combination and calibration. *Quart. J. Roy. Meteor. Soc.,* **135A,** 777–794, doi:10.1002/qj.383.

Kiladis, G. N., and K. M. Weickmann, 1992: Extratropical forcing of tropical Pacific convection during winter. *Mon. Wea. Rev.,* **120,** 1924–1938.

Klein, H., and Coauthors, 2010: Saharan dust and ice nuclei over central Europe. *Atmos. Chem. Phys.,* **10,** 10 211–10 221, doi:10.5194/acp-10-10211-2010.

Knippertz, P., 2005: Tropical–extratropical interactions associated with an Atlantic tropical plume and subtropical jet streak. *Mon. Wea. Rev.,* **133,** 2759–2776.

——, and J. Martin, 2005: Tropical plumes and precipitation in subtropical and tropical West Africa. *Quart. J. Roy. Meteor. Soc.,* **131,** 2337–2365.

——, and A. H. Fink, 2006: Synoptic and dynamic aspects of an extreme springtime Saharan dust outbreak. *Quart. J. Roy. Meteor. Soc.,* **132B,** 1153–1177, doi:10.1256/qj.05.109.

——, and J. E. Martin, 2007a: A Pacific moisture conveyor belt and its relationship to a significant precipitation event in the semiarid southwestern United States. *Wea. Forecasting,* **22,** 125–144.

——, and ——, 2007b: The role of dynamic and diabatic processes in the generation of cut-off lows over northwest Africa. *Meteor. Atmos. Phys.,* **96,** 3–19, doi:10.1007/s00703-006-0217-4.

Marticorena, B., and G. Bergametti, 1995: Modeling the atmospheric dust cycle: 1. Design of a soil-derived dust emission scheme. *J. Geophys. Res.,* **100,** 16 415–16 430.

Martin, J. E., 2006: *Mid-Latitude Atmospheric Dynamics: A First Course.* John Wiley and Sons, 336 pp.

Martius, O., E. Zenklusen, C. Schwierz, and H. C. Davies, 2006: Episodes of alpine heavy precipitation with an overlying elongated stratospheric intrusion: A climatology. *Int. J. Climatol.,* **26,** 1149–1164, doi:10.1002/joc.1295.

——, C. Schwierz, and H. C. Davies, 2010: Tropopause-level waveguides. *J. Atmos. Sci.,* **67,** 866–879.

Massacand, A. C., H. Wernli, and H. C. Davies, 1998: Heavy precipitation on the alpine southside: An upper-level precursor. *Geophys. Res. Lett.,* **25,** 1435–1438.

——, ——, and ——, 2001: Influence of upstream diabatic heating upon an alpine event of heavy precipitation. *Mon. Wea. Rev.,* **129,** 2822–2828.

Matsueda, M., 2009: Blocking predictability in operational medium-range ensemble forecasts. *SOLA,* **5,** 113–116, doi:10.2151/sola.2009–029.

McGuirk, J. P., A. H. Thompson, and N. R. Smith, 1987: Moisture bursts over the tropical Pacific Ocean. *Mon. Wea. Rev.,* **115,** 787–798.

——, ——, and J. R. Schaefer, 1988: An eastern Pacific tropical plume. *Mon. Wea. Rev.,* **116,** 2505–2521.

Palmer, T., F. Molteni, R. Mureau, R. Buizza, P. Chapelet, and J. Tribbia, 1992: Ensemble prediction. *Seminar on Validation of Models over Europe,* Vol. 1, ECMWF, 21–66. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom.]

Pappenberger, F., J. Bartholmes, J. Thielen, H. L. Cloke, R. Buizza, and A. de Roo, 2008: New dimensions in early flood warning across the globe using grand-ensemble weather predictions. *Geophys. Res. Lett.,* **35,** L10404, doi:10.1029/2008GL033837.

Park, Y., R. Buizza, and M. Leutbecher, 2008: TIGGE: Preliminary results on comparing and combining ensembles. *Quart. J. Roy. Meteor. Soc.,* **134,** 2029–2050, doi:10.1002/qj.334.

Richardson, D., R. Buizza, and R. Hagedorn, 2005: Final report of the 1st Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). WMO/TD-1273, WWRP/THORPEX Rep. 5, 39 pp.

Rossby, C.-G., 1959: Current problems in meteorology. *The Atmosphere and Sea in Motion,* B. Bolin, Ed., Rockefeller Institute Press, 9–50.

Schepanski, K., I. Tegen, B. Laurent, B. Heinold, and A. Macke, 2007: A new Saharan dust source activation frequency map derived from MSG-SEVIRI IR-channels. *Geophys. Res. Lett.,* **34,** L18803, doi:10.1029/2007GL030168.

Schwierz, C., S. Dirren, and H. C. Davies, 2004: Forced waves on a zonally aligned jet stream. *J. Atmos. Sci.,* **61,** 73–87.

——, P. Koellner-Heck, E. Z. Mutter, D. N. Bresch, P.-L. Vidale, M. Wild, and C. Schaer, 2010: Modelling European winter wind storm losses in current and future climate. *Climatic Change,* **101,** 485–514, doi:10.1007/s10584-009-9712-1.

Slingo, A., and Coauthors, 2006: Observations of the impact of a major Saharan dust storm on the atmospheric radiation balance. *Geophys. Res. Lett.,* **33,** L24817, doi:10.1029/2006GL027869.

Slingo, J. M., 1998: Extratropical forcing of tropical convection in a northern winter simulation with the UGAMP GCM. *Quart. J. Roy. Meteor. Soc.,* **124A,** 27–51.

Sodemann, H., A. S. Palmer, C. Schwierz, M. Schwikowski, and H. Wernli, 2006: The transport history of two Saharan dust events archived in an alpine ice core. *Atmos. Chem. Phys.,* **6,** 667–688.

Thorncroft, C. D., and H. A. Flocas, 1997: A case study of Saharan cyclogenesis. *Mon. Wea. Rev.,* **125,** 1147–1165.

——, B. J. Hoskins, and M. E. McIntyre, 1993: Two paradigms of baroclinic-wave life-cycle behaviour. *Quart. J. Roy. Meteor. Soc.,* **119,** 17–55.

Toth, Z., 1992: Quasi-stationary and transient periods in the Northern Hemisphere circulation series. *J. Climate,* **5,** 1235–1247.

——, and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74,** 2317–2330.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center—Practical aspects. *Wea. Forecasting,* **8,** 379–398.

Tripoli, G. J., C. M. Medaglia, S. Dietrich, A. Mugnai, G. Panegrossi, S. Pinori, and E. A. Smith, 2005: The 9–10 November 2001 Algerian flood: A numerical study. *Bull. Amer. Meteor. Soc.,* **86,** 1229–1235.

Wei, M., and Z. Toth, 2003: A new measure of ensemble performance: Perturbation versus error correlation analysis (PECA). *Mon. Wea. Rev.,* **131,** 1549–1565.

——, ——, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus,* **60A,** 62–79, doi:10.1111/j.1600-0870.2007.00273.x.

Wernli, H., and M. Sprenger, 2007: Identification and ERA-15 climatology of potential vorticity streamers and cutoffs near the extratropical tropopause. *J. Atmos. Sci.,* **64,** 1569–1586.

Ziehmann, C., 2001: Skill prediction of local weather forecasts based on the ECMWF ensemble. *Nonlinear Processes Geophys.,* **8,** 419–428.